forming a vector element, such that for each known biological fragment there is a respective vector element representing the number of times the respective representation of that known biological fragment is found in the subject protein sequence;

(c) from the formed vector elements, forming a vector having a length equal to the fixed number of known biological fragments in the provided set, such that the formed vector provides a uniform representation of the subject protein sequence; and

(d) providing the vector for making at least one analysis of the subject protein sequence, the uniform representation of the vector providing normalized input for the analysis.

20. (New) Apparatus for analyzing protein sequences, comprising:

a data store of a predefined number of known biological sequences; and

a comparison routine executed by a digital processor having access to the data store, the comparison routine comparing each known biological sequence from the data store to a subject protein sequence and generating a score indicative of the comparison, said scores forming a vector having a length equal to the predefined number of known biological sequences, such that said comparison routine provides the formed vector as a uniform representation of the subject protein sequence for at least one analysis, the uniform representation of the formed vector providing normalized input for the analysis.

B2 Cont

Please amend Claims 1, 3, and 10, 12 and 13. Amendments to the claims are indicated in the attached "Marked Up Version of Amendments" (pages i - iii).

---

1. (Amended) A method for analyzing a subject genome sequence comprising the steps of:

B3

providing a set of known biological fragments, the set being of a fixed number of said known biological fragments, each known biological fragment in the set having a respective representation;

comparing the respective representation of each known biological fragment from the set to a subject genome sequence, for each known biological fragment said comparing including (i) counting the number of times the respective representation of the known

biological fragment is found in the subject genome sequence and (ii) from said counted number of times, forming a vector element, such that for each known biological fragment there is a respective vector element representing the number of times the respective representation of that known biological fragment is found in the subject genome sequence;

from the formed vector elements, forming a vector having a length equal to the fixed number of known biological fragments in the provided set, such that the formed vector provides a uniform representation of the subject genome sequence; and

providing the formed vector for use as input to a desired analysis, the uniform representation provided by the formed vector enabling the formed vector to serve as normalized input.

B3 Cont'd

3. (Amended) A method as claimed in Claim 1 further comprising the step of:

B4

for each desired subject genome sequence, using said set of known biological fragments, repeating the comparing and forming steps such that a respective vector representation is formed and each desired subject genome sequence has a respective vector representation of a same length, said set of known biological fragments being a same set used for all of said subject genome sequences.

10. (Amended) Apparatus for analyzing genome sequences, comprising:

B5

a data store of a predefined number of known biological sequences;

a comparison routine executed by a digital processor having access to the data store, the comparison routine comparing each known biological sequence from the data store to a subject genome sequence and generating a score indicative of the comparison, said scores forming a vector having a length equal to the predefined number of known biological sequences, such that said comparison routine provides the formed vector as a uniform representation of the subject genome sequence and the formed vector enables at least one analysis of the subject genome sequence, the uniform representation of the formed vector providing normalized input for the analysis.

12. (Amended) Apparatus as claimed in Claim 10 further comprising a plurality of different subject genome sequences; and

wherein, using a same set of known biological sequences, the comparison routine forms, for each subject genome sequence, a respective vector such that a corresponding plurality of uniform length vector representations is provided.

76

13. (Amended) Apparatus as claimed in Claim 12 wherein the output of the comparison routine feeds the corresponding plurality of uniform length vector representations into further analysis processors.